Children's social referencing reflects sensitivity to graded uncertainty

Abstract

The ability to monitor epistemic uncertainty is critical for self-directed learning. However, we still know little about young children's ability to detect uncertainty in their knowledge. Here we asked whether a spontaneous information gathering behavior – social referencing – is related to epistemic uncertainty during early childhood. Children ages 2-5 completed a word-learning task in which they were presented with one or two objects, heard a label, and were asked to put the labeled object in a bucket. Referential ambiguity was manipulated through the number of objects present and their familiarity. In Experiment 1, when there were two novel objects and a novel label, the referent was ambiguous; when there were two familiar objects, or only one novel or familiar object, the referent was known or could be inferred. In Experiment 2, there were either two novel objects, two familiar objects, or one familiar and one novel object; in the latter case the referent could be inferred by excluding the familiar object. To further manipulate the availability of referential cues, the experimenter gazed at either the target or the center of the table while labeling the object. In both experiments, children looked at the experimenter more often while making their response when the referent was ambiguous. In Experiment 2, children also looked at the experimenter more when there was one familiar and one novel object, but only when the experimenter's gaze during labeling was uninformative. These results suggest that children's social referencing is a sensitive index of graded epistemic uncertainty.

Preschoolers quickly learn new concepts, rules, and language. They also actively explore (Schulz & Bonawitz, 2007) and ask questions (Chouinard, Harris, & Maratsos, 2007) in ways that seem targeted to maximize learning. These behaviors appear to require children to monitor epistemic states of ignorance and uncertainty, but young children have generally been credited with limited ability to monitor mental states (Sodian, Thoermer, Kristen, & Perst, 2012). To better understand how young children actively construct meaning from their environment, more evidence is needed about the relation between epistemic uncertainty and information seeking in early childhood. Do young children monitor uncertainty and engage in active learning behaviors based on this monitoring, or is early learning better characterized as a process of integrating information that is generated externally, for example, by social partners who act as teachers (Csibra & Gergely, 2006)?

Early metacognitive studies suggested that young children had limited capacities to reflect on their own thoughts or knowledge. For example, in one study (Markman, 1977), first-graders failed to recognize that they did not understand the instructions for a task unless they were prompted to act it out, while older children successfully identified the gaps in their understanding without prompting. In another study, 5-year-olds were found to have limited awareness of their own ongoing thoughts (J. H. Flavell, Green, Flavell, Harris, & Astington, 1995). However, these and other studies tended to rely on children's explicit verbal reports about their mental operations, often in a free-response format, which might underestimate children's ability to track epistemic states.

Recently, others have investigated children's uncertainty monitoring using tasks that do not require a verbal response from children. In one study, 3- to 5-year-olds completed perceptual discrimination and lexical identification tasks and reported on their certainty in their choices using a pictorial confidence scale (Lyons & Ghetti, 2011). Preschoolers reported being less confident when they responded incorrectly, suggesting that they were aware of their likelihood of accuracy based on their epistemic states. In another study, 3.5-year-olds completed a memory task in which they could opt out of responding to individual trials (Balcomb & Gerken, 2008). Children performed worse on trials they had opted out of when they were forced to answer them later on, suggesting that children used the opt-out option strategically to avoid answering when they were uncertain about their responses. On the other hand, Hembacher and Ghetti (2014) found that 3-year-olds reported being equally confident for correct and incorrect responses in a memory task using a pictorial confidence scale, while 4- and 5-year-olds were less confident when incorrect, suggesting that there are limits to 3-year-olds abilities to monitor uncertainty, and that some cognitive states may be more easily reflected on by young children than others. In sum, preschool-aged children demonstrate an emerging awareness of epistemic uncertainty in laboratory tasks that do not require a verbal or open-ended response. However, even

the competencies reported in these studies may underestimate children's ability to track epistemic uncertainty and act on it. The following section outlines an emerging body of research that suggests that children actively gather data to support learning from a young age, in a way that reflects sensitivity to epistemic uncertainty.

Active Learning in Early Childhood

First, there are studies that take inspiration from comparative metacognition research and ask whether children's spontaneous information-seeking behaviors track uncertainty. For example, Call and Carpenter (2001) had 2-year-olds choose between several tubes to find a hidden sticker. They found that the toddlers were more likely to peek inside a tube before choosing when they had not seen the baiting of the tubes compared to when they had, suggesting they were aware of their knowledge or ignorance and selectively sought confirmatory evidence before committing to a response when they did not know a sticker's location. In another study, Goupil, Romand-Monnier, and Kouider (2016) trained 20-month-olds to look at their parents if they needed help with a memory task in which they had to locate a hidden toy. Toddlers were more likely to seek help when they were completely ignorant of the toy's location, and when the task was more difficult because the delay between hiding and test was longer.

Others have focused on children's spontaneous exploration of objects under different epistemic conditions. Stahl & Feigenson (2015) found that 11-month-old infants were more likely to explore objects that had violated physical expectations (e.g., a toy car that appeared to hover in midair) than those that had made similar trajectories without appearing to violate physics (e.g., a toy that moved along a platform). Intriguingly, infants were more likely to explore the objects in ways that were suited to testing the physical property they had seen violated (e.g., picking up and dropping the car that had hovered in midair; banging a toy that had appeared to move through a solid wall). These results suggest that infants selectively explore objects to disambiguate causal properties that have been rendered uncertain by a surprising event.

There is also evidence that young children selectively explore when evidence is confounded (i.e., when existing evidence supports multiple causal hypotheses). For example, Schulz & Bonawitz (2007) found that preschoolers spent more time playing with a toy when they witnessed confounded evidence about its causal structure. Children who experienced a demonstration of a toy in which two levers were simultaneously depressed and caused two simultaneous events (a toy duck and puppet popping up) spent more time playing with that toy compared to children who saw a demonstration in which one lever was depressed at a time, revealing the unique function of each lever. This finding is consistent with the prediction that preschool-aged children track causal ambiguity and spontaneously explore more when they do not have sufficient evidence to isolate a cause for an effect.

Young children are also capable of independently generating the evidence they need to distinguish between causal hypotheses. Sim and Xu (2017) found that 2- and 3-year-olds were able to spontaneously discover first- and second-order rules to activate a toy (e.g., the triangle block makes the triangle toy work and toys are activated by shape-matched blocks) during free-play. Preschoolers also identify the most informative questions to confirm or disconfirm hypotheses (Ruggeri, Sim, & Xu, 2017). These behaviors would be unlikely to occur if children did not have some sensitivity to their epistemic states; they must track uncertainty in order to intervene effectively and gather disambiguating data.

These studies highlight that even young children are active learners who explore based on their epistemic states. They also suggest that spontaneous information-seeking behaviors may be a fruitful behavioral index of children's sensitivity to uncertainty. However, much of this research is confined to situations in which children must discover the causal properties of interesting objects. Does epistemic uncertainty drive other types of information seeking across learning domains? A great deal of early learning involves social partners; what is the role of uncertainty monitoring in children's social learning? In the next section we outline previous research on the selectivity of early social learning.

Selective social learning

Many current theories of early learning emphasize the role of social partners. Starting at birth, infants are particularly attuned to social information, preferring face-like stimuli over other configurations (M. C. Frank, Amso, & Johnson, 2014). Paying special attention to social partners allows infants and children to learn efficiently by witnessing others' behaviors and bypassing direct exploration or experience (Csibra & Gergely, 2006). Orienting towards social partners also allows infants and children to take advantage of direct pedagogy. Some have argued that early mindreading abilities afford infants and children the ability to interpret pedagogical information as such, allowing for stronger inferences from pedagogical compared to non-pedagogical information (Shafto, Goodman, & Frank, 2012).

Importantly, children are selective in their social learning. Children track the quality of informants' testimony, and prefer to learn from people who have a history of accurate testimony (Corriveau & Harris, 2009; Sobel & Kushnir, 2013). For example, older preschoolers choose to learn the words for novel objects from someone who has previously labeled objects correctly compared to incorrectly (Koenig, Clement, & Harris, 2009), and from people who appear confident rather than uncertain (Sabbagh & Baldwin, 2001). They also prefer to learn from people who have proven to be honest rather than dishonest in the past (Li, Heyman, Xu, & Lee, 2014). In general, young children appear to be selective learners with regard to the sources of social information they seek.

Are children also selective about *when* they seek social information? That is, do they seek social information only when their uncertainty about their own mental representations is high? There is some evidence to suggest that they do, at least in the context of solving problems of varying difficulty. For example, Vredenburgh & Kushnir (2015) found that preschool-aged children selectively sought help with a task when they were either less competent at the task or when they reached a difficult step. Coughlin, Hembacher, Lyons and Ghetti (2014) found that preschoolers were more likely to seek help on trials of a perceptual discrimination task for which they reported lower confidence in a separate session in which help was not available. Additionally, analysis of a corpus of parent-child conversations showed that preschoolers tend to ask questions that are directed to resolve gaps in knowledge, and they ask follow-up questions if they receive uninformative responses (Chouinard et al., 2007).

Help-seeking is only one form of social information-gathering. Infants and children also engage in social referencing, or checking a social partner's face for gaze direction and/or reactions to stimuli or events (Walden & Ogan, 1988). Young infants have been found to reference social partners to determine the safety of objects or actions; for example, infants reference their parents' faces before crossing a visual cliff of uncertain depth (Sorce, Emde, Campos, & Klinnert, 1985). Infants also check adults' faces when they are exposed to potentially dangerous stimuli or sounds, such as the sound of dogs barking or spiders (Striano & Rochat, 2000; Zarbatany & Lamb, 1985). Thus, infants and young children are prompted to reference social information by uncertainty about the safety of a situation.

Social referencing also plays a particularly important role in language acquisition. Referencing a social partner can provide several types of disambiguating information during word learning. For example, infants and children can follow a speaker's gaze direction to infer the referent of a new word, as people tend to look at the objects they refer to. By the second year of life infants follow a speaker's gaze and map labels to objects on the basis of gaze direction (Baldwin, 1991). There is also evidence that infants' propensity for gaze-following predicts later language development, highlighting the importance of this behavior for learning. In addition to monitoring gaze direction, children may reference a social partner's emotional reaction to a stimulus or event, which can help disambiguate the appropriateness of a response (Walden & Ogan, 1988). Finally, looking at a social partner can be taken as a bid for help (Vredenburgh & Kushnir, 2015), and may result in explicit instruction.

Social referencing can be an efficient source of disambiguating information, but is it driven by epistemic uncertainty during early childhood? It could be that social referencing is typically not costly enough to require selectivity, or that uncertainty signals related to knowledge representations are too weak to drive information-seeking behaviors in young children. Similarly, other learning mechanisms such as the privileging of social information (Ho, MacGlashan, Littman, & Cushman, 2017) or tracking of regularities in the environment

(Yurovsky & Frank, 2015) may be sufficiently powerful to obviate the need for uncertainty-dependent social referencing in preschool-aged children.

Current study

The present work asks whether preschoolers reference a speaker more frequently when the referent of the speech is ambiguous. This work adapts a paradigm used by Vaish, Demir and Baldwin (2011) in which 13- to 18-month-olds sat across from an experimenter who produced a label (e.g., "a modi!") in the presence of one or two novel objects. Infants looked towards the experimenter more often when there were two objects present, suggesting that infants' social referencing is driven by referential ambiguity. Here we adapt this procedure for use with preschoolers, who have a richer behavioral repertoire compared to infants, and may not reference social information based on uncertainty for the reasons discussed previously.

We ask whether preschoolers look more at a social partner when they are uncertain about the identity of a referent (Experiment 1) and whether they are sensitive to graded uncertainty based on the amount of disambiguating evidence available (Experiment 2). Furthermore, we examine social referencing across the time course of an event that includes a speaker labeling an object and requesting that the child put the object in a bucket. Thus, we can determine whether children differentially reference a speaker depending on the social information available at different moments. For example, children might reference a speaker during labeling to discover the gaze direction of the speaker, and during their choice about which object to pick to determine the speaker's evaluation of their choice.
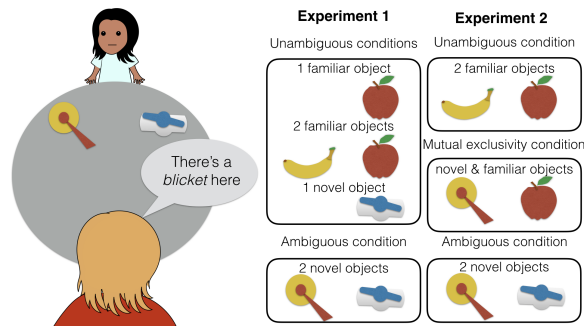

Figure 1: Study design for Experiments 1 and 2.

Experiment 1

In Experiment 1, we examined whether children would visually reference a speaker more often when the speaker produced a referentially ambiguous label compared to an unambiguous label. Children sat across from an experimenter who labeled an object on the table between them (Figure 1). The experimenter then asked the child to place the named object in a bucket. Across trials, there were either one or two objects on the table, which were either familiar or novel to the child. This design allowed us to test whether merely having more than one object present is sufficient to increase social referencing (which could not be ruled out by Vaish et al.), or if referential ambiguity (and thus epistemic uncertainty) is the underlying factor. If the latter is true, we expected children to increase their looking to the experimenter only on trials with two unfamiliar objects, when the object-label mapping was not known and could not be inferred.

We were interested in the amount of social referencing children exhibited across the trial. We considered four different phases of each trial based on the notion that children might expect different social information at different stages of the task. Specifically, we predicted that children might expect the speaker's gaze direction to be informative during the labeling itself, as speakers tend to look at objects they refer to. We predicted

that later in the trial, as children reached for an object and placed it in the bucket, they might expect evaluative feedback about their choice (e.g., facial expressions of encouragement or discouragement).

Methods

Participants

We recruited a planned sample of 80 children ages 2-5 years from the Children's Discovery Museum in San Jose, California.[1] The sample included 20 2-year-olds (mean age 31.97 months), 20 3-year-olds (mean age 42.65 months), 20 4-year-olds (mean age 55.85 months), and 20 5-year-olds (mean age 65.21 months). An additional 20 children participated but were removed from analyses because they heard English less than 75% of the time at home ($n = 10$), because they were unable to complete at least half of the trials in the task ($n = 4$), because of parental interference ($n = 1$), or due to experimenter or technical errors ($n = 5$).

Stimuli and Design

Children were presented with one or two objects, heard a label, and were asked to put the labeled object in a bucket. Half of the objects were selected to be familiar to children (e.g., a cow) and half were selected to be novel (e.g., a nozzle). There were four trial types: one-familiar, one-novel, two-familiar, and two-novel. There were three trials of each type, for a total of twelve trials. Trial types were presented sequentially in an order that was counterbalanced across participants. The assignment of individual objects to trial types was counterbalanced. On familiar trials, the familiar label for the target object was used (e.g., "cow"). On novel trials, a novel label was used (e.g., "dawnoo").

The critical manipulation was of referential ambiguity; one-familiar and two-familiar trials were referentially unambiguous, as children were expected to be certain about the objects and their labels. Similarly, on one-novel trials, children were expected to be certain about the label referent as there was only one option. However, two-novel trials were referentially ambiguous, as the novel label could apply to either novel object.

Throughout the task, the experimenter never gazed at the object they were labeling, or responded to children's verbal or non-verbal bids for help by indicating the correct object. Thus, children were expected to remain uncertain about the referent throughout the trial when two novel objects were present.

Procedure

Throughout the study, the child sat at one end of a large circular table, and the experimenter stood at the opposite end. Each trial of the task proceeded as follows: the experimenter placed one or two objects on the left and/or right sides of the table, out of reach of the child so that the child could not interact with the toys during the labeling event. For one-object trials, the location of the object (left or right) alternated between trials.

After placing the objects, the experimenter said "Hey look, there's a (target) here." The experimenter gazed at the center of the table rather than the object they labeled. The experimenter waited approximately two seconds based on a visual metronome placed within view before saying, "Can you put the (target) in the bucket?" They then pushed the object(s) forward within reach of the child, and placed a plastic bucket in the center of the table, also within reach of the child. Prior to the twelve experimental trials, there were two training trials: a one-familiar trial and a two-familiar trial, to acquaint the child with the procedure. A camera placed to the side of the experimenter captured the participant's face, so that looking behavior could be coded from video.

---

[1]Planned sample size, exclusion criteria, and analysis plan preregistered at https://osf.io/y7mvt

Coding procedure

Videos were coded using DataVyu software (http://datavyu.org). For each participant, we coded the number of times they referenced the experimenter across the trial. Because we were interested in the precise circumstances in which children feel uncertain enough to reference a speaker, we coded the number of looks that occurred during four phases of the trial: a *label* phase, which began at the utterance of the label and ended when the experimenter began to slide the objects, a *slide* phase, in which the experimenter slid the object(s) into the child's reach, a *planning* phase, which began at the end of the slide and ended when the child touched an object, and a *response* phase, which began when the child touched an object and ended when the child released the object into the bucket. A second coder independently scored the number of looks for one third of the trials for each participant to establish reliability.

Results and Discussion

Phase durations and accuracy

	Experiment	Trial Type	Gaze Condition	Age in Years	Mean Accuracy	CI lower bound	CI upper bound
1	1	2-fam	no gaze	2.00	0.77	0.71	0.82
2	1	2-fam	no gaze	3.00	0.97	0.95	0.99
3	1	2-fam	no gaze	4.00	0.92	0.89	0.96
4	1	2-fam	no gaze	5.00	0.99	0.97	1.00
5	2	2-fam	gaze	3.00	0.99	0.97	1.00
6	2	2-fam	gaze	4.00	0.94	0.91	0.96
7	2	2-fam	no gaze	3.00	0.95	0.92	0.97
8	2	2-fam	no gaze	4.00	1.00	1.00	1.00
9	2	mutual	gaze	3.00	0.97	0.95	0.99
10	2	mutual	gaze	4.00	1.00	1.00	1.00
11	2	mutual	no gaze	3.00	0.80	0.75	0.84
12	2	mutual	no gaze	4.00	0.96	0.94	0.98
13	2	2-nov	gaze	3.00	0.82	0.78	0.87
14	2	2-nov	gaze	4.00	0.84	0.80	0.88

Table 1: Accuracy Rates

	Experiment	Phase	Mean Duration	SD
1	1	label	5072.84	899.16
2	1	slide	872.43	196.83
3	1	planning	1253.85	2215.56
4	1	response	3646.40	4510.01
5	2	label	5239.23	748.05
6	2	slide	818.95	1307.27
7	2	planning	790.88	10327.45
8	2	response	4084.56	5032.37

Table 2: Phase Durations

Table 2 displays the average durations of each of the four phases. The *label* and *response* phases are longer on average than the *planning* and *slide* phases. However, note that we are interested in comparing the amount of social referencing across ambiguity conditions and not across phases directly.

We next examined children's accuracy (i.e., their rates of putting the correct item in the bucket for two-familiar trials, in which there was a correct response). Children sometimes put two items in the bucket (20.00% of trials), despite instructions to choose only one. If they put the correct item in the bucket first followed by the

incorrect object, we counted their response as correct. If they put the incorrect object in first, we counted their response as incorrect. Children also occasionally declined to choose an item (0.83% of trials); these trials are excluded from accuracy analyses. Children's accuracy is displayed in Table 1. Overall children generally chose the correct item for 2-familiar trials, indicating that they understood the task and were motivated to answer correctly. While 3- to 5-year-olds performed close to ceiling (92% - 99%), 2-year-olds were less accurate (76%), but still performed significantly above chance ($t = 43.20$, $p < .001$).
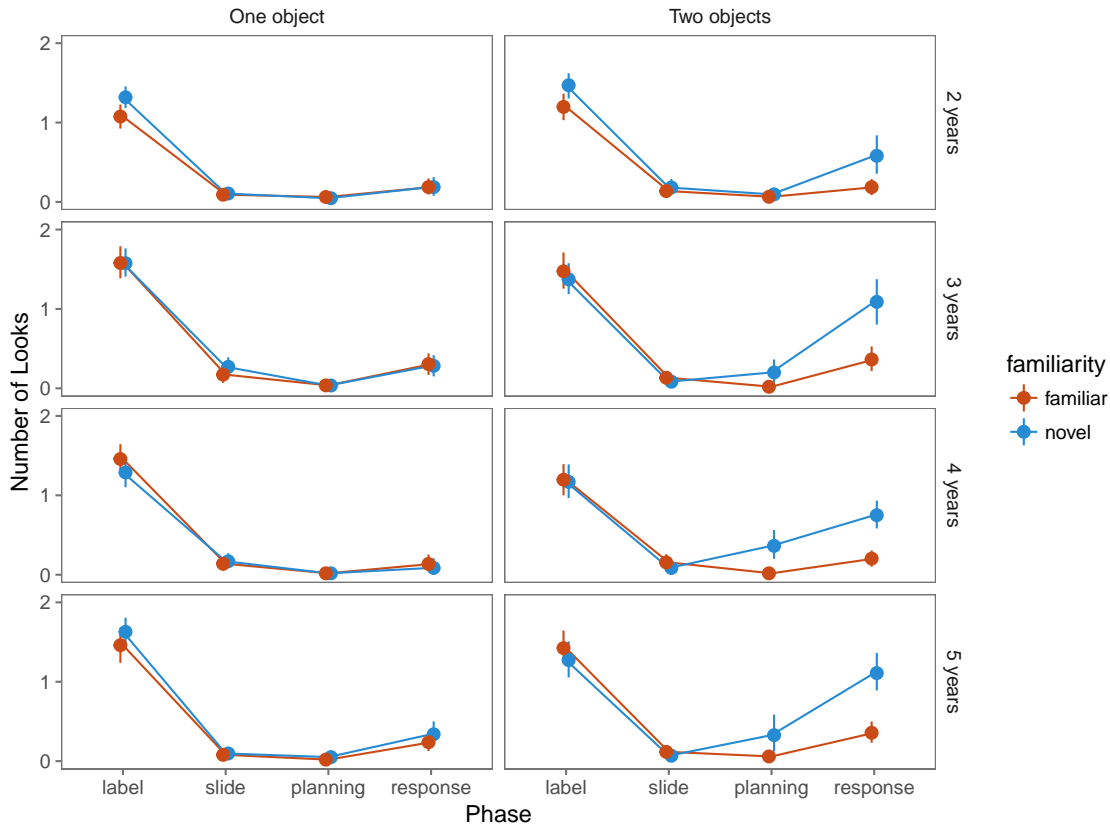

Figure 2: Results of Experiment 1. Number of looks to the experimenter across phases and conditions. Error bars are 95 percent confidence intervals. Age in months was entered as a continuous variable in regression models but is presented here as a categorical variable for visual simplicity.

Social referencing

Results of Experiment 1 are presented in Figure 2. Inter-rater reliability for the number of looks in each phase was high, intraclass correlation $r = .97$, $p<.001$. To test our prediction that referential ambiguity (i.e., having two novel objects) would produce more social referencing, we fit mixed-effects linear regression models separately for each phase with the following structure: `number of looks ~ number of objects * familiarity * age in months + (number of objects + familiarity | subject)`. A single model with phase as a factor did not converge.

We did not find any main or interaction effects of number of objects, familiarity, or age on number of looks during the *label* or *slide* phases. Thus, mere novelty or the presence of multiple objects was not enough to increase social referencing. However, we found an interaction effect of number of objects and familiarity during the *planning* ($\beta = 0.21$, $p < .001$) and *response* phases ($\beta = 0.60$, $p < .001$), such that 2-novel trials were associated with more looking. There was no interaction with age in either phase.[2] In summary, children

---

[2]https://github.com/emilyfae/socref_uncert

looked to the experimenter more often when planning and executing a response under uncertainty. These results suggest that children were aware that they did not have sufficient knowledge to answer independently, and referenced the speaker to resolve this uncertainty.

We did not find the expected effect of referential ambiguity in the *label* phase. It is possible that children failed to predict that they would need more information until later in the trial, when they were actually faced with making a decision. Another possibility is that children's looking was at ceiling during the labeling phase, perhaps because children tend to look at someone who is speaking regardless of the need for referential disambiguation. A third possibility is that this finding is an artifact of our design, in which the experimenter gazed at the center of the table rather than the referent of the label. Children may have realized that the experimenter's gaze direction during labeling was not informative. Alternatively, children may have found it strange to interact with an experimenter who did not gaze at the object they were labeling, which may have produced unnatural patterns of social referencing. Experiment 2 tests these possibilities and examines whether children's social referencing is sensitive to graded uncertainty.

Experiment 2

Experiment 2 was designed to replicate Experiment 1 and investigate whether children's social referencing is sensitive to uncertainty based on graded evidence about a label's referent. A hallmark of successful uncertainty monitoring is being less confident when the probability of accuracy is lower (Kiani & Shadlen, 2009). This ability includes awareness of complete ignorance, but also of graded evidence in mental representations (Fleming, Dolan, & Frith, 2012; Ghetti, Lyons, Lazzarin, & Cornoldi, 2008), which may be important for predicting outcomes and guiding behavior.

To examine children's social referencing on the basis of graded referential evidence, we added 1-novel-1-familiar trials. For these trials, we expected that children would be able to infer the referent by excluding the familiar object as a possibility. For example, when a toy lion and a novel item were present, they could exclude that the speaker was referring to the lion as a "blicket" (Markman & Wachtel, 1988). Since we did not observe any difference between one-familiar and one-novel trials, we eliminated single-object trials, leaving the 2-familiar and 2-novel trials. We predicted that children might be less certain about their choice on these trials compared to when the label and referent were familiar to them (2-familiar trials), but more confident than when there are no cues to reference (2-novel trials).

To further manipulate the availability of referential evidence, we varied between participants whether or not the experimenter's gaze during labeling was informative (they gazed at either the referent of their label or the center of the table). This manipulation also allowed us to determine whether children selectively reference gaze during labeling when gaze is informative. In Experiment 1, we did not observe an effect of age, so we restricted the current sample to 3- and 4-year-olds.

Methods

Participants

We recruited a planned sample of 80 children ages 3-4 years from the Children's Discovery Museum in San Jose, California.[3] The sample included 40 3-year-olds (mean age 42.89 months) and 40 4-year-olds (mean age 53.47 months). An additional 20 children participated but were removed from analyses because they heard English less than 75% of the time at home ($n = 9$), because they were unable to complete at least half of the trials in the task ($n = 7$), or due to experimenter or technical errors ($n = 4$).

---

[3]Planned sample size, exclusion criteria, and analysis plan preregistered at https://osf.io/y7mvt/.

Stimuli and Design

The stimuli and design were similar to Experiment 1 but included three trial types: 2-familiar, 2-novel, and 1-novel-1-familiar. There were four of each trial type, totaling twelve trials. In addition, we manipulated the experimenter's gaze behavior between participants. For half of the participants, the experimenter looked at the center of the table while labeling objects; for the remaining half, they looked directly at the objects they labeled.

Procedure

The procedure was identical to Experiment 1, except that there were three practice trials (two familiar trials and one novel trial). We included two familiar trials during the practice so that children would remain motivated to complete the task.

Results and Discussion

Phase durations and accuracy

The mean durations of the phases for Experiment 2 are presented in Table 2. They varied in length similar to Experiment 1. We additionally examined children's accuracy for familiar and mutual exclusivity trials, as well as novel trials in the Gaze condition, as all of these had a "correct" answer. Children's accuracy was calculated using the same criteria as in Experiment 1 (Table 1).

To quantify the effects of trial type, gaze condition, and age in months on the probability of accuracy, we fit the following mixed-effects logistic regression model: `correct ~ trial type * gaze + age + (1| subject)`. Accuracy was significantly lower for novel (β = -2.27, p < .001) and mutual trials (β = -2.23, p < .001) compared to familiar trials, and trial type interacted with gaze condition such that accuracy was significantly higher in the gaze condition for novel (β = 3.41, p < .001) and mutual trials (β = 1.12, p < .001). Age did not significantly predict accuracy (β = 0.18, p = .12).

Social referencing and referential ambiguity

Children's social referencing based on trial type and gaze condition are presented in Figure 3. Inter-rater reliability for the number of looks in each phase was again high, intraclass correlation r = .97, p<.001. To quantify the main and interactive effects of familiarity, gaze informativity, phase, and age on social referencing, we fit a mixed-effects linear regression model with the following structure: `number of looks ~ familiarity * age in months * gaze * phase + (familiarity | subject)`. In contrast to Experiment 1, a model with phase as a predictor converged.

First, do children reference a speaker more often when the objects and label are novel? Phase interacted with familiarity such that the *response* phase of novel trials was associated with significantly more looks (β = 0.28, p < .001). This result is consistent with results from Experiment 1. However, in contrast to Experiment 1, the number of looks was not significantly greater for novel trials in the *planning* phase.

We were also interested in whether mutual exclusivity trials would elicit an intermediate amount of uncertainty. We observed a three-way interaction of familiarity, gaze, and phase, such that the response phase of mutual exclusivity trials in the no-referential-gaze condition was associated with significantly more looks (β = -0.20, p < .01). Thus, mutual exclusivity trials were associated with greater looking only when the experimenter did not provide informative gaze. This finding is intriguing given that children should be able to solve mutual exclusivity trials without gaze information. Instead, these results suggest that children relatively uncertain while making a decision if excluding the familiar object is their only cue to reference, but this uncertainty is resolved if the speaker's gaze is informative. On the other hand, informative gaze during labeling did not lessen social referencing for novel trials, suggesting that gaze information alone was not sufficient to reduce

uncertainty. Instead, both gaze information and mutual exclusivity provided evidence about a label-object pairing, and children required both types of evidence to feel certain about their response.

Finally, we observed a four-way interaction such that the *response* phase of novel trials in the gaze condition was associated with more looking with increasing age (β = -0.02, $p <$ .01), suggesting that children may become more selective in their social referencing as they get older. It may be that children improve in their ability to monitor the need for disambiguating information, or they may become more likely to recognize that social information can be a source of disambiguation.

We did not observe selective social referencing during the *label* phase, even when referential gaze was available. This result rules out the possibility that children are less selective during labeling because they learned that gaze direction was not informative.
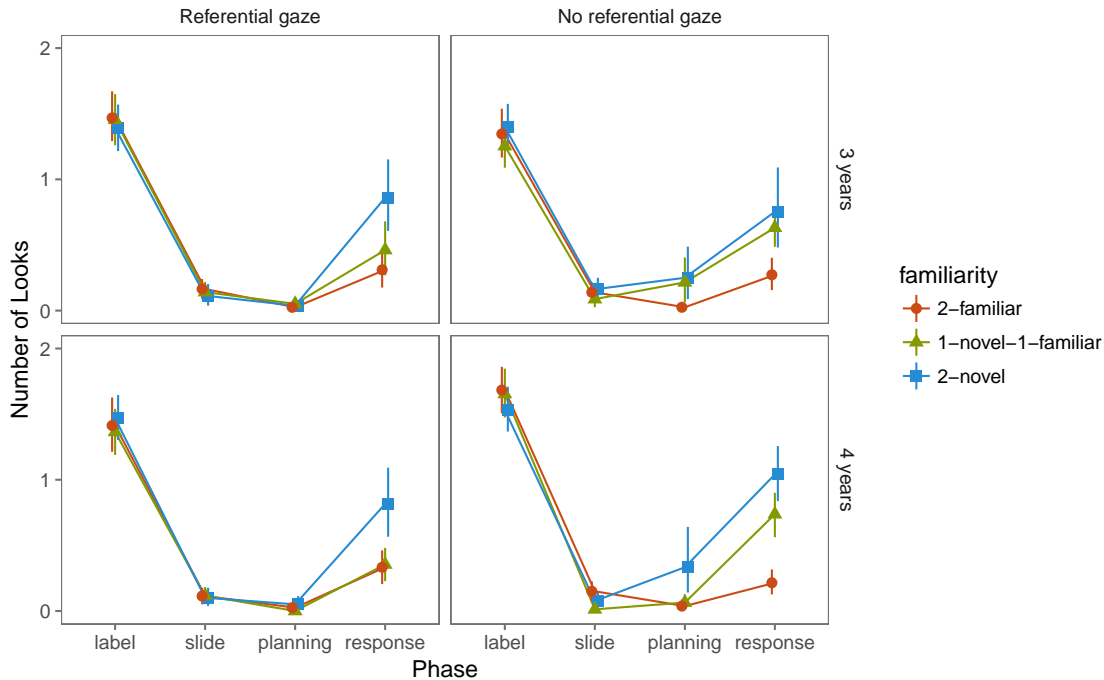

Figure 3: Results of Experiment 2. Number of looks to the experimenter across phases and trial types. Error bars are 95 percent confidence intervals.

Social referencing and accuracy

Another window into the relation between epistemic uncertainty and social referencing is to examine the amount of social referencing associated with accurate and inaccurate responses from children. Since incorrect responses are more likely to be associated with decision uncertainty (Kiani & Shadlen, 2009), more looks to the experimenter on incorrect trials would suggest that children reference social information when they are uncertain.

To quantify the relation between accuracy and social referencing across phases, we fit a mixed-effects linear regression model with the following structure: `number of looks ~ accuracy * age in months * phase + (1 | subject)`. Results showed that accuracy and phase interacted such that children looked at the experimenter significantly more during the labeling phase of accurate trials (β = 0.41, $p <$ .001), and significantly less in the *response* phase of accurate trials (β = -0.99, $p <$ .001).

In summary, children were more likely to look at the experimenter as they executed their response when they answered incorrectly, suggesting that social referencing is associated with epistemic uncertainty. This pattern increased with age. Additionally, children were more likely to answer correctly if they referenced the

speaker during the labeling phase. This pattern could be due to children obtaining referential information based on the experimenter's gaze in the gaze condition, or it could represent trials in which children were more engaged in the task.
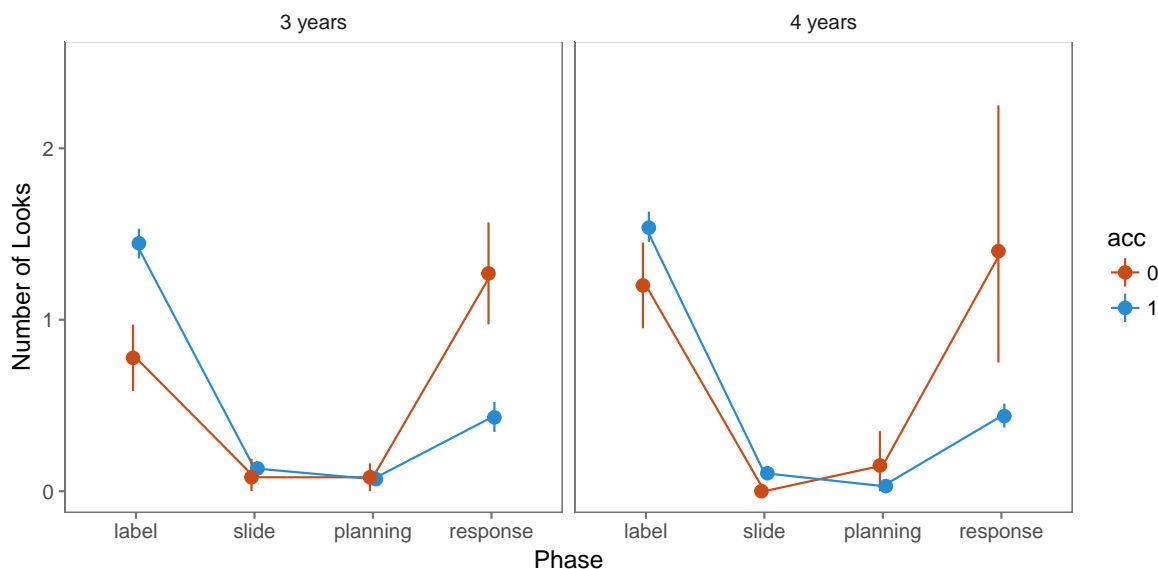

Figure 4: Results of Experiment 2. Number of looks to the experimenter based on accuracy. Error bars are 95 percent confidence intervals. Age in months was entered as a continuous variable in regression models but is presented here as a categorical variable for visual simplicity.

General Discussion

During the preschool years, children are increasingly able to actively gather information through help-seeking and exploration (Chouinard et al., 2007; Schulz & Bonawitz, 2007). Do children monitor their own uncertainty to guide these behaviors, or are they indiscriminate with regard to underlying knowledge states? Here, we examined whether young children's social referencing during a word-learning task was driven by uncertainty about a label's referent.

We found that referential ambiguity strongly predicted children's social referencing. Specifically, we observed this selectivity when children were forced to decide which object the speaker was referring to. We speculate that children referenced the speaker during the decision process because they expected evaluative feedback about their choice, either implicitly through the adult's facial expressions, or through an explicit response. This idea is consistent with other recent research that has found that preschoolers and toddlers seek help selectively when a problem is difficult or they are less skilled (Goupil et al., 2016; Vredenburgh & Kushnir, 2015).

Most intriguingly, we found that children's looking was driven by graded referential evidence. In the case of mutual exclusivity trials, children could solve the problem of reference by excluding the familiar item (Markman & Wachtel, 1988). Thus, unlike novel trials, they likely had some signal about the correct object-label mapping. If children simply monitored the presence or absence of such signals, they would have consistently treated mutual exclusivity trials as familiar trials. Instead, their social referencing depended on a combination of cues from mutual exclusivity and gaze informativity, suggesting that they are sensitive to graded evidence and seek disambiguating information only when uncertainty is relatively high. Children's greater social referencing on trials with only one cue to reference (i.e., mutual exclusivity trials with no referential gaze and novel trials with referential gaze) additionally suggests that children may remain uncertain about a new label-object mapping if they have not received confirmation of its accuracy, for example, through explicit feedback or gaze direction.

On the other hand, we found no evidence for selective social referencing as the object was being labeled. One possibility is that young children do not recognize the need for disambiguating information until they need to make a decision. Another possibility is that preschool-aged children spontaneously look at a speaker regardless of ambiguity, and additional looking was not needed or possible. Notably, Vaish et al. observed selective referencing during labeling among infants. Since infants in that study were holding one of the objects during labeling, referencing the speaker would have required them to disengage from that object, and may therefore have been more costly, promoting selectivity. Future research with preschoolers that includes a greater trade-off between attentional options would help to distinguish among these possibilities. Overall, these results provide evidence that preschool-aged children monitor graded uncertainty in their mental representations and act on that uncertainty through spontaneous information-seeking.

Acknowledgements

References

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*(5), 750–760.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–890.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, *3*(4), 207–220.

Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*, 1–129.

Corriveau, K., & Harris, P. L. (2009). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Developmental Science*, *12*(1), 188–193.

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2014). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, *18*(6), 957–971.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development* (pp. 249–274). Oxford: Oxford University Press: Oxford University Press.

Flavell, J. H., Green, F. L., Flavell, E. R., Harris, P. L., & Astington, J. W. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development*, i–iii–v–vi–1–113.

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1280–1286.

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*(C), 13–26.

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *Journal of Experimental Child Psychology*, *99*(3), 157–181.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they dont

know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492–3496.

Hembacher, E., & Ghetti, S. (2014). Dont look at my answer: Subjective uncertainty underlies preschoolers exclusion of their least accurate memories. *Psychological Science*, *25*(9), 0956797614542273–1776.

Ho, M. K., MacGlashan, J., Littman, M. L., & Cushman, F. (2017). Social is special: A normative framework for teaching with and learning from evaluative feedback. *Cognition*, 1–16.

Kiani, R., & Shadlen, M. N. (2009). Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex. *Science*, *324*, 759–764.

Koenig, M. A., Clement, F., & Harris, P. L. (2009). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *15*, 694–698.

Li, Q.-G., Heyman, G. D., Xu, F., & Lee, K. (2014). Young children's use of honesty as a basis for selective trust. *Journal of Experimental Child Psychology*, *117*(C), 59–72.

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, *48*(3), 986–992.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121–157.

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Science*.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers theory of mind and semantic development. *Child Development*, *72*, 1054–1070.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351.

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from two- and three-year-old children. *Developmental Psychology*.

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*, 779=797.

Sodian, B., Thoermer, C., Kristen, S., & Perst, H. (2012). Metacognition in infants and young children. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of metacognition* (pp. 119–133).

Sorce, J. F., Emde, R. N., Campos, J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, *21*, 195–200.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants learning and exploration. *Science*, *348*(6230), 91–94.

Striano, T., & Rochat, P. (2000). Emergence of selective social referencing in infancy. *Infancy*, *1*, 253–264.

Vaish, A., Demir, Ö. E., & Baldwin, D. (2011). Thirteen- and 18-month-old infants recognize when they need referential information. *Social Development*, *20*(3), 431–449.

Vredenburgh, C., & Kushnir, T. (2015). Young children's help-seeking as active information gathering. *Cognitive Science*, *40*(3), 697–722.

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*(5), 1230–1240.

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning.

Cognition, 145(C), 53–62.

Zarbatany, L., & Lamb, M. E. (1985). Social referencing as a function of information source: Mothers versus strangers. *Infant Behavior and Development, 8*, 25–33.